OIPE

ENTERED

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/942,087A

DATE: 05/09/2002
TIME: 13:00:36

Input Set : A:\seqlist.txt
Output Set: N:\CRF3\05092002\I942087A.raw

```
 3 <110> APPLICANT: Case, Casey Christopher
 4        Wolffe, Alan
 5        Urnov, Fyodor
 6        Lai, Albert
 7        Snowden, Andrew
 8        Tan, Siyuan
 9        Gregory, Philip
11 <120> TITLE OF INVENTION: MODULATION OF ENDOGENOUS GENE EXPRESSION IN CELLS
13 <130> FILE REFERENCE: 8325-0002.21 / S2-US5
15 <140> CURRENT APPLICATION NUMBER: 09/942,087A
16 <141> CURRENT FILING DATE: 2001-08-28
18 <150> PRIOR APPLICATION NUMBER: 09/229,037
19 <151> PRIOR FILING DATE: 1999-01-12
21 <160> NUMBER OF SEQ ID NOS: 43
23 <170> SOFTWARE: PatentIn Ver. 2.0
25 <210> SEQ ID NO: 1
26 <211> LENGTH: 25
27 <212> TYPE: PRT
28 <213> ORGANISM: Artificial Sequence
30 <220> FEATURE:
31 <223> OTHER INFORMATION: Description of Artificial Sequence:exemplary motif
32        of C2H2 class of zinc finger proteins (ZFP)
34 <220> FEATURE:
35 <221> NAME/KEY: MOD_RES
36 <222> LOCATION: (2)..(3)
37 <223> OTHER INFORMATION: Xaa = any amino acid
39 <220> FEATURE:
40 <221> NAME/KEY: MOD_RES
41 <222> LOCATION: (4)..(5)
42 <223> OTHER INFORMATION: Xaa = any amino acid, may be present or absent
44 <220> FEATURE:
45 <221> NAME/KEY: MOD_RES
46 <222> LOCATION: (7)..(18)
47 <223> OTHER INFORMATION: Xaa = any amino acid
49 <220> FEATURE:
50 <221> NAME/KEY: MOD_RES
51 <222> LOCATION: (20)..(22)
52 <223> OTHER INFORMATION: Xaa = any amino acid
54 <220> FEATURE:
55 <221> NAME/KEY: MOD_RES
56 <222> LOCATION: (23)..(24)
57 <223> OTHER INFORMATION: Xaa = any amino acid, may be present or absent
59 <400> SEQUENCE: 1
```

RAW SEQUENCE LISTING   DATE: 05/09/2002
PATENT APPLICATION: US/09/942,087A   TIME: 13:00:36

Input Set : **A:\seqlist.txt**
Output Set: **N:\CRF3\05092002\I942087A.raw**

```
W--> 60 Cys Xaa Xaa Xaa Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
     61    1               5                  10                  15
W--> 63 Xaa Xaa His Xaa Xaa Xaa Xaa Xaa His
     64              20                  25
     65 <210> SEQ ID NO: 2
     66 <211> LENGTH: 10
     67 <212> TYPE: DNA
     68 <213> ORGANISM: Artificial Sequence
     70 <220> FEATURE:
     71 <223> OTHER INFORMATION: Description of Artificial Sequence:ZFP target site
     72       with two overlapping D-able subsites
     74 <220> FEATURE:
     75 <221> NAME/KEY: modified_base
     76 <222> LOCATION: (1)..(2)
     77 <223> OTHER INFORMATION: n = g,a,c or t
     79 <220> FEATURE:
     80 <221> NAME/KEY: modified_base
     81 <222> LOCATION: (5)
     82 <223> OTHER INFORMATION: n = g,a,c or t
     84 <220> FEATURE:
     85 <221> NAME/KEY: modified_base
     86 <222> LOCATION: (8)
     87 <223> OTHER INFORMATION: n = g,a,c or t
     89 <220> FEATURE:
     90 <221> NAME/KEY: modified_base
     91 <222> LOCATION: (9)
     92 <223> OTHER INFORMATION: n = a,c or t; if g, then position 10 cannot be g
     93       or t
     95 <220> FEATURE:
     96 <221> NAME/KEY: modified_base
     97 <222> LOCATION: (10)
     98 <223> OTHER INFORMATION: n = a or c; if g or t, then position 9 cannot be g
    100 <400> SEQUENCE: 2
W-->101 nngkngknnn                                                10
    104 <210> SEQ ID NO: 3
    105 <211> LENGTH: 10
    106 <212> TYPE: DNA
    107 <213> ORGANISM: Artificial Sequence
    109 <220> FEATURE:
    110 <223> OTHER INFORMATION: Description of Artificial Sequence:ZFP target site
    111       with three overlapping D-able subsites
    113 <220> FEATURE:
    114 <221> NAME/KEY: modified_base
    115 <222> LOCATION: (1)..(2)
    116 <223> OTHER INFORMATION: n = g,a,c or t
    118 <220> FEATURE:
    119 <221> NAME/KEY: modified_base
    120 <222> LOCATION: (5)
    121 <223> OTHER INFORMATION: n = g,a,c or t
```

RAW SEQUENCE LISTING   DATE: 05/09/2002
PATENT APPLICATION: US/09/942,087A   TIME: 13:00:36

Input Set : A:\seqlist.txt
Output Set: N:\CRF3\05092002\I942087A.raw

```
        124 <220> FEATURE:
        125 <221> NAME/KEY: modified_base
        126 <222> LOCATION: (8)
        127 <223> OTHER INFORMATION: n = g,a,c or t
        129 <400> SEQUENCE: 3
W--> 130 nngkngkngk                                              10
        133 <210> SEQ ID NO: 4
        134 <211> LENGTH: 5
        135 <212> TYPE: PRT
        136 <213> ORGANISM: Artificial Sequence
        138 <220> FEATURE:
        139 <223> OTHER INFORMATION: Description of Artificial Sequence:linker
        141 <400> SEQUENCE: 4
        142 Asp Gly Gly Gly Ser
        143   1               5
        146 <210> SEQ ID NO: 5
        147 <211> LENGTH: 5
        148 <212> TYPE: PRT
        149 <213> ORGANISM: Artificial Sequence
        151 <220> FEATURE:
        152 <223> OTHER INFORMATION: Description of Artificial Sequence:linker
        154 <400> SEQUENCE: 5
        155 Thr Gly Glu Lys Pro
        156   1               5
        159 <210> SEQ ID NO: 6
        160 <211> LENGTH: 9
        161 <212> TYPE: PRT
        162 <213> ORGANISM: Artificial Sequence
        164 <220> FEATURE:
        165 <223> OTHER INFORMATION: Description of Artificial Sequence:linker
        167 <400> SEQUENCE: 6
        168 Leu Arg Gln Lys Asp Gly Glu Arg Pro
        169   1               5
        172 <210> SEQ ID NO: 7
        173 <211> LENGTH: 4
        174 <212> TYPE: PRT
        175 <213> ORGANISM: Artificial Sequence
        177 <220> FEATURE:
        178 <223> OTHER INFORMATION: Description of Artificial Sequence:linker
        180 <400> SEQUENCE: 7
        181 Gly Gly Arg Arg
        182   1
        185 <210> SEQ ID NO: 8
        186 <211> LENGTH: 5
        187 <212> TYPE: PRT
        188 <213> ORGANISM: Artificial Sequence
        190 <220> FEATURE:
        191 <223> OTHER INFORMATION: Description of Artificial Sequence:linker
        193 <400> SEQUENCE: 8
```

RAW SEQUENCE LISTING   DATE: 05/09/2002
PATENT APPLICATION: US/09/942,087A   TIME: 13:00:36

Input Set : A:\seqlist.txt
Output Set: N:\CRF3\05092002\I942087A.raw

194 Gly Gly Gly Gly Ser
195   1   5
198 <210> SEQ ID NO: 9
199 <211> LENGTH: 8
200 <212> TYPE: PRT
201 <213> ORGANISM: Artificial Sequence
203 <220> FEATURE:
204 <223> OTHER INFORMATION: Description of Artificial Sequence:linker
206 <400> SEQUENCE: 9
207 Gly Gly Arg Arg Gly Gly Gly Ser
208   1   5
211 <210> SEQ ID NO: 10
212 <211> LENGTH: 9
213 <212> TYPE: PRT
214 <213> ORGANISM: Artificial Sequence
216 <220> FEATURE:
217 <223> OTHER INFORMATION: Description of Artificial Sequence:linker
219 <400> SEQUENCE: 10
220 Leu Arg Gln Arg Asp Gly Glu Arg Pro
221   1   5
224 <210> SEQ ID NO: 11
225 <211> LENGTH: 12
226 <212> TYPE: PRT
227 <213> ORGANISM: Artificial Sequence
229 <220> FEATURE:
230 <223> OTHER INFORMATION: Description of Artificial Sequence:linker
232 <400> SEQUENCE: 11
233 Leu Arg Gln Lys Asp Gly Gly Gly Ser Glu Arg Pro
234   1   5   10
237 <210> SEQ ID NO: 12
238 <211> LENGTH: 16
239 <212> TYPE: PRT
240 <213> ORGANISM: Artificial Sequence
242 <220> FEATURE:
243 <223> OTHER INFORMATION: Description of Artificial Sequence:linker
246 <400> SEQUENCE: 12
247 Leu Arg Gln Lys Asp Gly Gly Gly Ser Gly Gly Gly Ser Glu Arg Pro
248   1   5   10   15
251 <210> SEQ ID NO: 13
252 <211> LENGTH: 25
253 <212> TYPE: DNA
254 <213> ORGANISM: Artificial Sequence
256 <220> FEATURE:
257 <223> OTHER INFORMATION: Description of Artificial Sequence:ZFP target site
258   region surrounding initiation site of vascular
259   endothelial growth factor (VEGF) gene containing
260   two 9-base pair target sites
262 <220> FEATURE:
263 <221> NAME/KEY: protein_bind

RAW SEQUENCE LISTING   DATE: 05/09/2002
PATENT APPLICATION: US/09/942,087A   TIME: 13:00:36

Input Set : A:\seqlist.txt
Output Set: N:\CRF3\05092002\I942087A.raw

```
264 <222> LOCATION: (4)..(12)
265 <223> OTHER INFORMATION: upstream 9-base pair ZFP VEGF1 target site
267 <220> FEATURE:
268 <221> NAME/KEY: protein_bind
269 <222> LOCATION: (14)..(22)
270 <223> OTHER INFORMATION: downstream 9-base pair ZFP VEGF3a target site
272 <400> SEQUENCE: 13
273 agcgggagg atcgcggagg cttgg                                        25
276 <210> SEQ ID NO: 14
277 <211> LENGTH: 298
278 <212> TYPE: DNA
279 <213> ORGANISM: Artificial Sequence
281 <220> FEATURE:
282 <223> OTHER INFORMATION: Description of Artificial Sequence:VEGF1 ZFP
283        construct targeting upstream 9-base pair target
284        site in VEGF promoter
286 <220> FEATURE:
287 <221> NAME/KEY: CDS
288 <222> LOCATION: (2)..(298)
289 <223> OTHER INFORMATION: VEGF1
291 <400> SEQUENCE: 14
292 g gta ccc ata cct ggc aag aag aag cag cac atc tgc cac atc cag ggc 49
293   Val Pro Ile Pro Gly Lys Lys Lys Gln His Ile Cys His Ile Gln Gly
294   1               5                   10                  15
296 tgt ggt aaa gtt tac ggc aca acc tca aat ctg cgt cgt cac ctg cgc  97
297 Cys Gly Lys Val Tyr Gly Thr Thr Ser Asn Leu Arg Arg His Leu Arg
298             20                  25                  30
300 tgg cac acc ggc gag agg cct ttc atg tgt acc tgg tcc tac tgt ggt 145
301 Trp His Thr Gly Glu Arg Pro Phe Met Cys Thr Trp Ser Tyr Cys Gly
302         35                  40                  45
304 aaa cgc ttc acc cgt tcg tca aac ctg cag cgt cac aag cgt acc cac 193
305 Lys Arg Phe Thr Arg Ser Ser Asn Leu Gln Arg His Lys Arg Thr His
306     50                  55                  60
307 acc ggt gag aag aaa ttt gct tgc ccg gag tgt ccg aag cgc ttc atg 241
308 Thr Gly Glu Lys Lys Phe Ala Cys Pro Glu Cys Pro Lys Arg Phe Met
309  65                  70                  75                  80
311 cgt agt gac cac ctg tcc cgt cac atc aag acc cac cag aat aag aag 289
312 Arg Ser Asp His Leu Ser Arg His Ile Lys Thr His Gln Asn Lys Lys
313                 85                  90                  95
315 ggt gga tcc                                                      298
316 Gly Gly Ser
319 <210> SEQ ID NO: 15
320 <211> LENGTH: 99
321 <212> TYPE: PRT
322 <213> ORGANISM: Artificial Sequence
324 <220> FEATURE:
325 <223> OTHER INFORMATION: Description of Artificial Sequence:VEGF1 ZFP
326        construct targeting upstream 9-base pair target
327        site in VEGF promoter
```

RAW SEQUENCE LISTING ERROR SUMMARY   DATE: 05/09/2002
PATENT APPLICATION: US/09/942,087A   TIME: 13:00:37

Input Set : A:\seqlist.txt
Output Set: N:\CRF3\05092002\I942087A.raw

Please Note:
Use of n and/or Xaa have been detected in the Sequence Listing. Please review the
Sequence Listing to ensure that a corresponding explanation is presented in the <220>
to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:1; Xaa Pos. 2,3,4,5,7,8,9,10,11,12,13,14,15,16,17,18,20,21,22,23,24
Seq#:2; N Pos. 1,2,5,8,9,10
Seq#:3; N Pos. 1,2,5,8
Seq#:41; N Pos. 1,2
Seq#:42; N Pos. 1,2
Seq#:43; N Pos. 1,2

VARIABLE LOCATION SUMMARY   DATE: 05/09/2002
PATENT APPLICATION: US/09/942,087A   TIME: 13:00:37

Input Set : A:\seqlist.txt
Output Set: N:\CRF3\05092002\I942087A.raw

<u>Use of n's or Xaa's(NEW RULES):</u>
Use of n's and/or Xaa's have been detected in the Sequence Listing.
Use of <220> to <223> is MANDATORY if n's or Xaa's are present.
in <220> to <223> section, please explain location of n or Xaa, and which
residue n or Xaa represents.

Seq#:1; Xaa Pos. 2,3,4,5,7,8,9,10,11,12,13,14,15,16,17,18,20,21,22,23,24
Seq#:2; N Pos. 1,2,5,8,9,10
Seq#:3; N Pos. 1,2,5,8
Seq#:41; N Pos. 1,2
Seq#:42; N Pos. 1,2
Seq#:43; N Pos. 1,2

VERIFICATION SUMMARY   DATE: 05/09/2002
PATENT APPLICATION: **US/09/942,087A**   TIME: 13:00:37

Input Set : **A:\seqlist.txt**
Output Set: **N:\CRF3\05092002\I942087A.raw**

L:60 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1 after pos.:0
L:63 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1 after pos.:16
L:101 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2 after pos.:0
L:130 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3 after pos.:0
L:865 M:258 W: Mandatory Feature missing, <221> not found for SEQ ID#:41
L:865 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:41
L:865 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:41 after pos.:0
L:877 M:258 W: Mandatory Feature missing, <221> not found for SEQ ID#:42
L:877 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:42
L:877 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:42 after pos.:0
L:889 M:258 W: Mandatory Feature missing, <221> not found for SEQ ID#:43
L:889 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:43
L:889 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:43 after pos.:0